DOCUMENT RESUME

ED 380 748 CG 026 164

AUTHOR Powell, Richard
TITLE Comparing Methods of Evaluating Treatment Effects.
PUB DATE [95]
NOTE 15p.
PUB TYPE Information Analyses (070)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Comparative Analysis; Counseling; *Counseling Effectiveness; *Counselor Evaluation; *Counselor Performance; *Evaluation Criteria; Evaluation Methods; Evaluation Needs; *Evaluation Problems; Outcomes of Treatment; Psychologists; Therapy

ABSTRACT

This paper examines four common measures of treatment outcomes--ethical standards, statistical analysis, standards of care, and social validation--and explores the inherent problems associated with each. Ethical standards may assure the public that only reasoned professional approaches are used, but the ethical treatment of patients does not mean that they experienced a beneficial change. Statistical analysis, on the other hand, can detect changes, but its primary use in group treatment offers little help to practitioners dealing with individual clients. A standard of care--where all practitioners approach a problem in a certain way--suffers from contradictions in treatment protocol since explicit standards are not available for many clinical problems. One suggested solution is to adopt an established series of steps in providing services. Lastly, social validation, a process which determines how society views the whole spectrum of treatment, faces problems similar to standards of care since normative rates for all clinically significant behaviors are not readily available. The best measure of outcome, it is argued, would use these four indices in concert. By combining measures, counselors can arrive at a more accurate picture of therapeutic effect and begin the process for wide agreement on measuring outcomes. (RJM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 380 748

Comparing Methods of Evaluating Treatment Effects

Richard Powell

Ohio University, School of Applied Behavioral Science and

Education Leadership

Running Head: Treatment Outcomes

CG026164

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

R. POWELL

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

BEST COPY AVAILABLE

Comparing Methods of Evaluating Treatment Effects

Evaluation of treatment outcome is required in this age of concern regarding health care reform. (Walz, 1990) Dollars available for health care, including mental health, are becoming scarce. The public and third party payers are demanding the best treatment for their money. We must then decide which treatments are providing help for the consumer and which are not.

Several measures can be used to evaluate treatment. Four such indices are ethical standards, statistical analysis, standards of care and social validation. Each of these measures have their strong points and limitations. The purpose of this paper is to examine these common indices of treatment effectiveness and inherent problems associated with each. Hopefully this will aid counselors and agencies in understanding the treatment evaluation options open to them.

Ethical Standards

Evaluation of how well treatment follows established ethical and legal guidelines is a way of assuring the public that only reasoned professional approaches were used. Kitchener (1994) delineated five principles by which ethical comparisons can be made. We can evaluate treatment along these guidelines which included autonomy. Autonomy means the person is treated with dignity and respect including given the opportunity to give informed consent to treatment. This idea is embodied in many of our social institutions including the practice of psychotherapy. The second principle is nonmaleficence. This principle requires that the treatment not harm the individual. The third principle is beneficence. This principle says that in order for the treatment to be considered good under ethical guidelines, it must produce some benefit for the person. Justice is the fourth principle. Justice requires that treatment must be provided in a just or fair fashion. The final principle is fidelity which requires the practitioner to perform as expected in the clients best interests.

Treatment can be evaluated on the basis of these five principles. We can easily see by reviewing consent forms if a person has been treated with respect to their autonomy. An absence of complaints from consumers is at best a partial indication that no harm was done and we have satisfied the principle of nonmalificence. We can review service delivery to make sure that treatment has not been used unfairly or in a discriminatory fashion. For instance, the service is offered to all consumers, in ways that show a sensitivity to affirmative action issues. This would satisfy the principle of justice. The evaluation of beneficence can present a number of unique difficulties. For example, the client may have come to therapy experiencing Post Traumatic Stress Disorder (American Psychiatric Association, 1994). At the conclusion of therapy the person still manifests all the important clinical features of the disorder. He or she are experiencing flashbacks, panic attacks, disruption in interpersonal functioning, etc., but has quit smoking because of the therapy. Smoking is harmful and so we could say the person had benefitted from therapy and satisfy the principle of beneficence. Further analysis may indicate that the therapist followed through faithfully in providing treatment and satisfied the principle of fidelity. As all five principles were satisfied we would be forced to conclude that the treatment was ethical but it would not be a significant change for the individual. It would appear then that ethical guidelines are not sufficient alone to evaluate treatment effects.

Statistical Techniques

Statistical analysis of important clinical behaviors is another technique useful to the practitioner for evaluation of important clinical behaviors . The practitioner will measure some important clinical behavior. Post treatment measures can be compared to pre treatment rates or

the rates of untreated individuals to evaluate change. Differences in observations are obtained by the practitioner and these differences are submitted to analysis of statistical tests. If the differences are large (statistically significant) the practitioner can feel that the differences are not due to chance and may then be the result of the applied treatment. (Spence, Underwood, Duncan, and Cotton, 1990) This is the standard for research and many consider it the standard for evaluating treatment data. (Barrios, 1990)

Statistical analysis of treatment data , does have serious limitations. Practitioners who use behavior therapy often point out that the effect of treatment on groups of individuals is not meaningful to practitioners dealing with the individual client. The practitioner is not concerned with how particular treatments affect individuals on average, but how a manipulation will affect the specific individual they are seeing. (Hersen, 1990)

Another obvious limitation of statistical analysis is seen in the head banging example by Kazdin. (1982) It is not uncommon for head banging to occur 100 times per day. Treatment may reduce this to 50 times a day. Such a large difference would surely be significant under any statistical procedure and we could feel confident that the results were due to our intervention rather than chance. Such a result would be significant statistically but it would not be significant clinically. The only acceptable rate for such behavior would be zero which should be readily apparent for any behavior likely to cause injury. In subtle problematic behaviors such as negative self talk, (Meichenbaum, 1977) or expressions of sadness, guilt and gloom (Beck, 1987) the acceptable post treatment rate may not be as obvious. Statistical analysis is a valid measure of treatment effects, but, alone it also is not a sufficient measure of effectiveness.

We have seen that ethical considerations and statistical analysis are valid measures of treatment effects. But neither alone are sufficient measures. There is some dimension not measured. Using both ethics and statistics in concert in our previous example will show that more is still needed. For example, one would surely admit that reducing head banging by 50 percent is beneficial to the individual, but still not enough. We must then look to an additional measure.

Standards of Care

Standard of care represent how treatment is typically done by trained professionals in certain situations. That is we look at how practitioners commonly approach a clinical situation. A standard of care is established when all practitioners approach the problem in a certain fashion.

Locating professional counseling standards of care are not easy since experts in counseling do not agree on how particular problems should be addressed. There is no manual of standards. So explicit standards are not available for many clinical problems. Sources could include continuing education activities (conferences and workshops) and the professional literature. What is presented here could be worth emulating and thus a standard of care. For instance, one standard of care for treatment plans is provided by Seligman (1990). She recommends that the practitioner approach service delivery from a comprehensive point of view. Using her acronym D.O. A. C.L.I.E.N.T. M.A.P., the practitioner proceeds through a series of steps in providing services. The steps form her acronym are as follows:

Diagnosis-clients diagnosis

Objectives of treatment

Assessments needed-such as physical or neurological

Clinician characteristics viewed as therapeutic

Location of services (for example in hospital or out patient)

Intervention to be used

Emphasis of treatment (level of directiveness, cognitive or affective emphasis)

Nature of treatment (individual, family, group)

Timing (frequency of sessions duration)

Medications needed

Adjunct services required (such as support groups, education)

Prognosis (Seligman, 1990)

Following this standard will provide competent, quality service. This standard is flexible and can be employed in a broad variety of clinical problems. This standard of care is readily available to most practitioners. We can look at treatment and see if a practitioner applied treatment or diagnosis within this standard. In fact this approach is used presently to evaluate treatment in the legal arena of malpractice. Not delivering services within the accepted standards may result in financial and legal penalties for the practitioner. To avoid malpractice, the practitioner need not deliver outstanding service, just service that would be normal or average for a practitioner in the geographic region. It is obvious then that the standard of care to avoid malpractice is not going to be the zenith measure of treatment effectiveness.

Unfortunately standards of care are not readily available for all client problems. In the case of malpractice, this standard is usually made available through the testimony of other practitioners. Practitioners with expertise in the problem area testify as to what is normal and this sets the standard. No catalogue for standards of care exists but there are calls for one to be developed in the future.

Another problem with relying solely on standards of care is that they are descriptions of how the problems are usually dealt with and do not cover new or exceptional cases. They leave little room for innovation which places the practitioner who departs from the standard at a significant risk.

It may not be possible from reviewing the professional literature to evaluate the effectiveness of different types of treatments. Smith, Glass, and Miller (1980) made a comparison of different treatment approaches and found that certain types of therapy produced better results than others. On the other hand, they could not see which specific aspects of these treatments made them more effective than the other treatment models. It is impossible to replicate treatment entirely from the treatment literature or to differentiate accurately between the effectiveness of different approaches. This problem occurs because all relevant treatment information may not be published. For instance, information is withheld to protect the client's confidentiality and this data may be needed to replicate treatment effectively. Treatment approaches also have many commonalities making differential evaluation ambiguous or difficult. (Stiles, Shapiro, and Elliot, 1986) Thus, even careful review of treatment literature does not always provide useful information for use by the practitioner.

We have seen then that standards of care are not readily available or easily discernible. This then limits the utility of these standards for evaluating treatment now.

It is important though for the practitioner to conform with standards of care in providing services. With this n mind the practitioner can consult with peers, review of the treatment literature, or attend workshops. Specifically this must be done to avoid malpractice. However as we have seen standards of care are not well defined now for most treatments and thus cannot be readily used in evaluating treatment. Also there are difficulties in using this measure with innovative treatments or in difficult cases. Like ethical guidelines and statistical analysis, standards of care are not sufficient individual measures of treatment.

Social Validation

Wolfe (1978) introduced a technique for measuring the clinical importance of treatment outcome. This was called social validation and while it has been used extensively in behavior therapy and applied behavior analysis it can be effectively used with any treatment approach. (Kazdin, 1983) This process involves generally determining how society views the focus of treatment, the procedures that are used, and the effect these procedures have.

Kazdin identified two methods for evaluating clinical importance of treatment effects. The first, social comparison, involves comparing the behavior of the client with the behavior of others. Those selected for comparison must be considered as non deviant and functioning adequately in their daily environment. The focus of treatment is then to make the problematic behavior similar to the targeted peers behavior. A simple example is provided from the classic study done by O'Brien and Azrin in 1972. The focus of treatment was the eating behavior of

retarded individuals. Before intervention the practitioners observed customers in a local restaurant. Here, the eating behavior of these customers was recorded and rated as the retarded individuals were to be scored in treatment. They found that the retarded individuals before treatment did significantly worse than the restaurant customers. After treatment the eating error rate of the clients was as good or better than the restaurant customers. It is important to note here that the rates of eating errors for restaurant customers and treated individuals never reached zero. If this had been the targeted goal this training may be continuing today. Social validation can therefore provide a technique of evaluating treatment effects and also help in developing reasonable goals for treatment.

The second method of social validation described by Kazdin (1983) is subjective evaluation. This is done by having others evaluate the client's behavior and judge its acceptability. These are typically individuals familiar with the person and with whom the person normally interacts. They are asked to view the effects of treatment either live or through video tapes and asked whether the change is meaningful. It is subjective in that the criterion for acceptability may not be known by the practitioner or the evaluator. In child therapy the individual is referred by parents who feel that the current state of affairs is unacceptable. In most instances, they would also be available and appropriate judges of post treatment effects. A difficulty in using this approach is finding valid and appropriate evaluators for adults. One source for evaluators for a specific client would be persons identified by the client. Behavior typically occurs in a social context and most clients come into therapy because their behavior is distressing to others. For adults, the client also may be the evaluator when impairment of

judgment is not a major issue. The chief thing to keep in mind in using this technique is to respect the confidentiality of the client.

We have seen that we may use two techniques to evaluate the importance of treatment effects. Social comparison involves the evaluating of the treatment change against normative samples. The second type of social validation involves having appropriate others view the treatment change and rate the significance of improvement. This method is called subjective evaluation. The two methods could be used in concert. For example, the subjective raters could be asked to view the treated person and persons not having problems. They then could be asked to objectively rate these two samples. These differences in behavior could be objectively identified and subsequently used as a treatment goals and objective measures of treatment effects.

There are difficulties in using social validation now. The first, like standards of care, normative rates for all clinical significant behavior are not readily available. But, they are more accessible than these standards. For instance, as we saw in the previous examples normative data may be available in a local area such as a restaurant or similar public arena. Another source is the Diagnostic and Statistical Manual of the American Psychiatric Association. Here in the diagnostic criteria are many objective standards for evaluating treatment. Many sources for these standards can be found but the practitioner may have to exert some effort or use some creativity. Also more professional journals could require measures of social validity for publication. This is common in the behavior therapy literature and as we said could be done with other approaches. This would then give us greater access to normative data in this regards.

care it can be done and not compromise this client right. Social validation therefore is an important technique for evaluating treatment effects.

We have seen that treatment effects can be evaluated in four ways. First, treatment can be evaluated by examining the compliance of the treatment with legal and ethical standards. Next we can employ the rigors of research and evaluate the outcome through statistical analysis. We can also evaluate treatment against a standard of care. A common practice particularly in avoiding malpractice. Finally, a technique from the behavioral therapies can be employed to evaluate the clinical importance of treatment change.

All the techniques identified in this paper have problems and limitations. These difficulties appear to arise when the technique is used alone as a measure. Possibly the best model would be to use the measures in concert. For example, we could use social validation and ethical evaluation together to measure a treatment outcome. This would appear to solve the problem with our example of the person with P.T.S.D. who following treatment had stopped smoking. Again we may find the treatment ethical, but submitting this outcome to social validation would reveal that the outcome is severely lacking.

We could also return to our example of the head banger. Again a reduction of the rate by one half may be statistically significant. We could also examine our treatment and find it ethical. Again analysis through social validation would point out the deficiencies.

Social validation also is not an answer. Used alone, without the other measures, it could cause a type of social tyranny we would definitely find unacceptable. Again the measures must be used in concert to be truly effective at determining quality of treatment.

This paper examined four common measures of treatment outcome and proposed as a model that all measures be used in concert. Without such an approach we as a profession cannot progress and will constantly disagree as to what constitutes good treatment. Like the four blind men trying to describe an elephant based on the part each felt, we are certain from one perspective that the treatment we use is successful or that the treatment someone else uses is not. We need as a profession to agree how we will measure outcome. Combining the measures, we will get a more accurate picture and one that will lead to agreement within the profession.

One fact is upon us. We must evaluate treatment effects. Eyesenk in 1966 found that psychotherapy had little effect. While this study was found in error, many in the public are skeptical regarding the effects of psychotherapy. Currently many insurance clerks feel that they can evaluate the effects of psychotherapy better than any professional and until we can put together a satisfactory model to do this they are probably right.

REFERENCES

American Psychiatric Association, Diagnostic And Statistical Manual-Fourth Edition American Psychiatric Association, Washington D.C., 1994.

Barrios, B.A. Experimental Design in Group Outcome Research. in Bellack, A.S., Hersen, M, and Kazdin, A.E., International Handbook of Behavior Modification and Therapy, Plenum Press, New York, 1990, 151-171.

Beck, A.T. Cognitive Therapy. In Zeig, J.K., The Evolution of Psychotherapy, Bruner/Mazel, New York, 1987, 149-178.

Eyesenk, H.J. The Effects of Psychotherapy. New York: International Science Press, 1966.

Hersen, M. Single Case Experimental Designs, in Bellack, A.S., Hersen, M, and Kazdin, A.E., International Handbook of Behavior Modification and Therapy, Plenum Press, New York, 1990, 151-171.

Kazdin, A.E. Methodological Strategies in Behavior Therapy Research, In Wilson, G.T. and Franks, C.M., Contemporary Behavior Therapy, Guilford Press, New York, 1982.

Kitchener, K.S. Intuition, Critical Evaluation, and Ethical Principles: The Foundation for Ethical Decisions in Counseling Psychology. The Counseling Psychologist, 12(3), 1984, 43-55.

Miechenbaum, D. Cognitive Behavior Modification: An Integrative Approach, Plenum, New York, 1977.

OBrien, F. and Azrin, N. Developing proper mealtime behaviors of the institutionalized retarded. Journal of Applied Behavior Analysis, 1972, 5, 389-399.

Seligman, L. Selecting Effective Treatments. Josey-Bass, San Francisco, 1990.

Smith, M.L., Glass, G.V., and Miller, T.I., The benefits of Psychotherapy. John Hopkins University Press, Baltimore, 1980.

Stiles, W.B., Shapiro, D.A., and Elliot, R., Are all Psychotherapies Equivalent?, American Psychologist, 1986, 41, 165-180.

Spence, J.T., Underwood, B.J., Duncan, C.P., and Cotton, J.W. Elementary Statistics, Appleton Century Crofts, New York, 1990

Walz, G.R. Focused Generalizations on Counseling, In Walz, G.R., Gazda, G.M. and Shertzer, B. Counseling Futures, Eric, 1991, 71-86.

Wolf, M.M. Social Validity: The Case For Subjective Measurement, or How Applied Behavior Analysis is Finding it Heart. Journal of Applied Behavior Analysis, 1978, 203-214.